



Spracherkennung Status und Ausblick

Daniel Sobe, 04.12.2021



Motivation

- Wie soll die Spracherkennungstechnologie eingesetzt werden?
- Anwendungen:
 - Sprachassistent / smart home (prototyp)
 - Spielzeug (Produkt)
 - Diktierfunktion (Dienst)
 - Simultanübersetzer / Automatische Untertitel (Dienst)
 - Weitere?



Anfänge: Digitalisierungskonferenz 2019

- Wenig Sprachressourcen → „klassisches“ System
 - Akustisches Modell (aufgezeichnete Sprache → Phonemkette)
 - Wörterbuch (Phoneme → Wortkette)
 - Sprachmodell (Wörter → erlaubte Wortgruppen)
- benötigt werden:
 - transkribierte Aufzeichnungen für akustisches Modell
 - geschriebene Texte
 - phonetisches / sprachwissenschaftliches Wissen



Ressourcen und Hilfe

- Aufzeichnungen im „Mozilla Common Voice“ Projekt
- Kooperation mit Fraunhofer IKTS und BTU Cottbus für Obersorbisch:
 - Statistische Auswertung von Textkorpora
 - Aufstellen der Phonemliste und der Ausspracheregeln
 - Auswahl Aufzeichnungen aus „Common Voice“ / Neue Aufzeichnungen
- Erfolgreiche Anwendung unter Zuhilfenahme eines deutschen akustischen Modells in einer Studie:
 - http://www.essv.de/pdf/2021_43_50.pdf



Ergebnisse der Studie

- Angepasstes akustisches Modell eignet sich gut für eine eingeschränkte Domäne (smart home)
- Demonstrator unter Open Source Lizenz veröffentlicht:
 - https://github.com/ZalozbaDev/web_demo_spoznawanje
- Prototyp „digidom“ unter Open Source Lizenz veröffentlicht:
 - https://github.com/ZalozbaDev/raspberry_pi_demo_spoznawanje
- Weitere Arbeit an Robustheit und Leistungsfähigkeit nötig



Fortführung 2021 - heute

- Akustisches Modell:
 - Übergang auf ein obersorbisches Modell
 - Sprecheranpassung und Anpassung an den akustischen Pfad
- Wörterbuch:
 - Verbesserung der Ausspracheregeln
- Sprachmodell:
 - Technologieverbesserung für die Erstellung großer Modelle
- Aufzeichnungen:
 - Anleitung zur Satzauswahl und Aufnahmebedingungen



Ausblick

- Veröffentlichung aller Ergebnisse aus 2021
- Niedersorbisch
- Geplante Kooperationen:
 - Sorbisches Institut: Phonetik / Linguistik für Ober- und Niedersorbisch
 - RCW: Kombination Spracherkennung mit [sotra.app](#)
 - FHG und BTU: Anwendungen mit großem Sprachmodell
- Hilfe bei der persönlichen Benutzung und Anpassung des „digidom“-Prototyps
- Weitere Aufzeichnungen



Ende / Fragen